PACT/Hazardexpert. However, by the rule they claim to have used (equivocal carcinogens are regarded as noncarcinogens), only γ-butyrolactone is predicted correctly.

Inclusion of additional related variables in the predictive method. After the carcinogenicity outcomes were known, Lewis et al. found that the predictive performance of COMPACT could be enhanced if they included an additional COMPACT prediction (C2E) and also a predictor variable ("Hazardexpert") that incorporates information about metabolism. While there is nothing inherently wrong with including additional variables in a predictive methodology, this exercise should ideally have been carried out prospectively, not retrospectively. It is much easier to find predictive variables that work once the study outcomes to be predicted are known. The important (and yet to be answered) question is, how will the authors' newly derived, multivariate predictive methodology fare for prospective predictions? Hopefully, it will be better than COMPACT's limited predictive success (56%) for the 44 NTP chemicals.

The combination of COMPACT and Hazardexpert eliminated the apparent discordance for three chemicals: tris(2-chloroethyl)phosphate, 2,3-dibromo-1-propanol, and 1,2,3-trichloropropane, while introducing discordance for another chemical previously predicted correctly (methyl bromide). However, the authors misclassify two other chemicals: chloramine and HC Yellow 4, both of which are reported as successful predictions, but in fact were not predicted correctly (see Table 2). Including additional predictor variables (and not correcting for the misclassification of chloramine and HC Yellow 4) reduced the number of discordant predictions from 11 to 8.

Inclusion of additional, apparently unrelated, variables in the predictive method. The eight chemicals that Lewis et al. conclude are not correctly predicted by COMPACT/Hazardexpert are designated in their Table 4 (p. 182). The authors then carry out further analyses to reduce the number of discordant predictions from eight to five. Frankly, it is unclear exactly how the authors achieve this reduction. It appears that the basis for eliminating the final three chemicals from "discordancy" was an appeal to "structural alert, chronic toxicity studies, and the Ames test," which correctly predicted the carcinogenicity of *o*-benzyl-*p*-chlorophenol, methylphenidate hydrochloride, and diphenylhydantoin, three chemicals "missed" by COMPACT/Hazardexpert. One other chemical (mercuric chloride) not even evaluated by COMPACT/Hazardexpert, but correctly identified by "the metal ion redox potentials for inorganic compounds," was also apparently added in as a correct prediction. The authors should justify how these additional predictions, based on apparently unrelated variables, can be meaningfully interpreted as improving the performance of COMPACT/ Hazardexpert. In any case, the authors include these successful predictions in their calculations and conclude that the concordance for COMPACT/ Hazardexpert when predicting rodent carcinogenicity is 86% (32/37). I leave it to the reader's judgment to determine how much confidence to place in this figure.

I have no objection to the development of techniques designed to predict rodent (or more importantly, human) carcinogenicity, and I suspect that it is possible to develop methods that will be successful in this regard. However, I strongly urge caution in placing too much confidence in COMPACT or in any other predictive method that has little success when applied prospectively and seems to work only when applied retrospectively to the original data set, using extensive (and scientifically questionable) data manipulations and reanalysis.

Joseph K. Haseman
National Institute of Environmental Health Sciences
Research Triangle Park, North Carolina

Reference

1. Lewis DFV, Ioannides C, Parke DV. A prospective toxicity evaluation (COMPACT) on 40 chemicals currently being tested by the National Toxicology Program. Mutagenesis 5:433–435 (1990).

Response

In response to Joe Haseman's letter, we would like to point out that although our article is retrospective with regard to the rodent carcinogenicity study of the 40 chemicals, the COMPACT data were available at the time of the release of the carcinogenicity assays. The Hazardexpert evaluations for the 40 chemicals were carried out after the NIEHS conference, but the Hazardexpert system (available commercially from Compudrug Ltd) is not part of COMPACT. The following account, hopefully, provides some clarification of the points raised in Dr. Haseman's letter.

Most other systems publish their predictions or analyses without providing any mathematical derivation which can be reproduced by others. In contrast, we show how our predictions/analyses are generated from numerical values (COMPACT parameters) for molecular and electronic features of each chemical. Our attempts to provide a numerical description of the COMPACT plot of molecular planarity/potential chemical reactivity, have not been entirely successful, due to the fact that the training set of chemicals shows a curved line discriminating P4501 specificity from other P450 isozymes, whereas the COMPACT ratio (either area/depth^2/ΔE or area/depth^2/ΔE – 8) gives a straight line relationship. This results in some chemicals (e.g., resorcinol) having a COMPACT ratio and a COMPACT graphical plot which give conflicting results, but the graph is the *original* paradigm. We have recently derived an expression which is more complex (*1*), based on analysis of the COMPACT curve, and this gives more precise results in terms of correlation with the graph, although the actual graphical representation is preferred.

Resorcinol was predicted to be positive in COMPACT using the COMPACT ratio, but the graph of area/depth^2 versus ΔE as presented at the 1993 NTP conference (*2*) clearly shows that this compound should be negative as it is outside the curve. This is the only example in all of the 40 chemicals of a discrepancy between the approximation of the COMPACT ratio and the accurate description of the graph. As the *EHP* paper is retrospective, we feel justified in making this point, even though the graphical description was available in the conference documentation. HC Yellow 4 was changed from positive in COMPACT to negative, due to the fact that the original structure sent to us by NTP was erroneous and was subsequently changed by NTP after our original predictions had been published. When we ran the new (correct) structure through our system, it proved negative, and we feel justified in making this clear in our retrospective study published in *EHP*. However, we provided revised data (including the aforementioned cases) and distributed this at the NTP conference, which, moreover, included our results for the P4502E descriptor, now provided in the February 1995 issue of *EHP* (103:178–184).

The Hazardexpert analyses were generated retrospectively as we had only recently purchased the software. As can be seen from our *EHP* paper, the Hazardexpert results (which utilize the EPA database) give quite good concordances with positive carcinogens, and they are better than the Ames test for negatives and also overall.

Regarding the relatively poor performance of the original computer-based predictions compared with that of Ashby, Tennant, and others, it should be emphasized that the latter employed a combination of mutagenicity, subchronic toxicity, and structural alert tests, which are, therefore, three evaluations combined into one prediction—so it is perhaps not surprising

that this combination of three different tests gives the best overall concordance with rodent carcinogenicity. It is well known that the Ames test gives only just over 50% concordance with rodent carcinogenicity, but it is still extensively used. This is because it has a well-defined endpoint that one can readily understand in biological terms, i.e., genotoxicity. However, there is overwhelming evidence to show that enzymes of the cytochrome P450 superfamily are involved in the metabolism and toxicity of most (~90%) chemicals. P450s play a pivotal role in toxicity and carcinogenicity, and the COMPACT system is designed to identify P450-mediated metabolism and metabolic activation. Although this system (*3*) was originally based solely on the structures of known P450 substrates, we have now generated full three-dimensional structures (*4*) of the mammalian enzymes themselves (including human isoforms) which agree closely with experimental findings. However, we are also aware that there are other mechanisms of carcinogenicity that do not require P450 activation, and structure alert systems can be useful in identifying direct-acting carcinogens, for example (*1,5*).

Nongenotoxic carcinogens are not so easy to predict but we are elaborating models to identify chemicals involved in peroxisome proliferation and other activation pathways such as ß-lyase cleavage. Eventually there will be a battery of tests in place, which we hope will adequately assess the likely risk to *man* from exposure to foreign compounds; so models of human enzymes and receptors which may mediate potentially carcinogenic events will be important. What is crucial is the determination of how readily a chemical is metabolized and whether any reactive intermediates (ROS or metabolites) are sufficiently long-lived to cause irreparable DNA damage. There may well be short-term test procedures developed which can assess these factors *in vitro*, but computer-based systems can be just as accurately predictive; however, these computer systems do not require synthesis of the chemical, are extremely rapid, and, consequently, relatively inexpensive. We appreciate that traditional toxicologists may have been suspicious of replacing biological tests with computer predictions, but there is evidence that attitudes are changing.

One reason for publishing our retrospective study of the 40 NTP chemicals was to show that it is possible for a combination of tests to give reasonable concordances with rodent carcinogenicity, and we did not anticipate that moderating the equivocal results of the rodent study in the light of the pathology report presented at the NTP conference would be controversial. Our *EHP* paper was independently refereed and, as our use of modified equivocal results (reported at the meeting) was not questioned by the referees, one can only presume that NIEHS (which publishes *EHP* and also organized the 1993 conference) did not find this contentious. The problem regarding equivocals is well known, and these are obviously difficult to assess by predictive systems which, in general, do not equivocate. In fact, some systems (e.g., TOPKAT, CASE) tend to exclude equivocal results in the rodent assay when they validate their methods. At the NIEHS conference in 1993 there was lengthy discussion about equivocals, and the prevailing doctrine from NTP was to regard these as negatives, although there was not universal agreement for this view among the delegates. During the presentation of the pathology results it was indicated that a few of the equivocals could be interpreted, on histopathological evidence, as being weakly positive. It did not seem unreasonable to us in our retrospective analysis to take into account the views of the NIEHS pathologist who conducted the examinations. However, if one excludes the equivocals, the concordance between COMPACT and the rodent assay becomes 70% (21/30), which is not much different from the concordance one gets from regarding three of the equivocals as positives. If one regards all of the equivocals as positives, the concordance between COMPACT and the rodent carcinogenicity is 69% (25/36), whereas taking them as negative lowers the concordance further to 64% (23/36). However, it should be noted that if these three equivocals are regarded as positive, most (if not all) of the predictive tests show a similar improvement.

The use of metal ion redox potentials for providing some estimate of carcinogenicity is not currently part of COMPACT, but it is of interest to show that physicochemical parameters may be employed to try to predict the potential carcinogenicity of inorganic compounds. Likewise, Hazardexpert is not part of COMPACT, but the two tests are fairly complementary in the comparisons we have made to date (*1,6*).

In our *EHP* paper we provide explanations of the results for each chemical, including possible reasons why some of these were discordant with the rodent assay. However, we have not altered our results in the light of this additional biochemical knowledge, especially when the predictive methods (e.g., Ames test, Ashby structural alert, etc.) were able to predict correctly the outcome of the rodent carcinogenicity study for those chemicals. The purpose of these developments of predictive systems is to be able to decrease the number of animal experiments and the time required for the safety evaluation of chemicals that are destined for human exposure, and we believe that a combination of several systems represents the best way to achieve this (*6*), with the consideration of P450-mediated pathways of activation and detoxication being most important. Our *EHP* paper merely attempts to show how a combination of systems might work, but we would appreciate advice from a statistician on how to "weight" such tests: perhaps Joe Haseman could help us.

D. F. V. Lewis
C. Ioannides
D. V. Parke
University of Surrey
Guildford, Surrey, UK

References

1. Brown SJ, Raja AA, Lewis DFV. A comparison between COMPACT and Hazardexpert evaluations for 80 chemicals tested by the NTP/NCI rodent bioassay. Alter Lab Anim 22:482–500 (1994).
2. Lewis DFV, Ioannides C, Parke DV. Computer-optimised molecular parametic analysis of chemical toxicity (COMPACT). In: Proceedings of the conference on predicting chemical carcinogenesis in rodents, 24–25 May 1992, Research Triangle Park, North Carolina. Research Triangle Park, NC:National Institute of Environmental Health Sciences, 1992; 45–49.
3. Lewis DFV, Ioannides C, Parke DV. Validation of a novel molecular orbital approach (COMPACT) to the safety evaluation of chemicals by comparison with Salmonella mutagenicity and rodent carcinogenicity data evaluated by the US NCI/NTP. Mutat Res 291:61–77 (1993).
4. Lewis DFV, Moereels H, Lake BG, Ioannides, Parke CD. Molecular modelling of enzymes and receptors involved in carcinogenesis: QSARS and COMPACT-3D. Drug Metab Rev 26:261–285 (1994).
5. Lewis DFV. Computer-assisted methods in the evaluation of chemical toxicity. Rev Computat Chem 3:173–222 (1992)
6. Lewis DFV. Comparison between rodent carcinogenicity test results of 44 chemicals and a number of predictive systems. Regul Toxicol Pharmacol 20:215–222 (1994).

Drinking Water and Leukemia

Cohn et al. recently expanded (*EHP* 102:556–561) on an earlier ecological study (*1*) of leukemia and drinking water in northern New Jersey. The initial study suggested an association between volatile organic hydrocarbon (VOC) contamination of drinking water and increased risk of leukemia among females (but not males). The authors concluded that 1) the appearance of an association for females only